



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/671,683	09/28/2000	Kazuyuki Ashimura	FUR0012-US	7247

28970 7590 04/07/2004

SHAW PITTMAN
IP GROUP
1650 TYSONS BOULEVARD
SUITE 1300
MCLEAN, VA 22102

EXAMINER

STORM, DONALD L

ART UNIT	PAPER NUMBER
----------	--------------

2654

DATE MAILED: 04/07/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/671,683

Applicant(s)

ASHIMURA ET AL.

Examiner

Donald L. Storm

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on September 28, 2000 through June 23, 2003.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-21 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-9, 13-16 and 21 is/are rejected.
- 7) ☒ Claim(s) 10-12 and 17-20 is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 28 September 2000 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date <u>6</u> . | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Oath/Declaration

1. The substitute oath or declaration submitted by the Applicant was received on December 13, 2000 (paper 3); it has been entered.

Specification

2. The title is objected to because it is not sufficiently descriptive of the invention. A new title is required that is clearly indicative of the invention to which the claims are directed. See MPEP § 606.01. The Examiner suggests that the Applicant consider a title including these elements: "Speech Synthesis Devices Handling Phoneme Units of Extended CV."

3. The abstract is objected to under 37 C. F. R. § 1.72 because it does not describe the disclosure sufficiently, particularly the matter claimed as new. A cursory inspection of the abstract should inform readers of the nature and gist of the technical disclosure. See MPEP § 608.01(b). Appropriate correction is required. The following additional matter contained in the disclosure should be briefly mentioned:

- a. Extended CV refers to a contiguous sequence of phonemes without clear distinction containing at least one vowel (claim 5 and others);
- b. contour of vocal tract transmission function of phonemes (claim 8).

Revision of the abstract's content may be needed to add the claimed subject matter and conform to the general guideline limit of 150 words. For example, the abstract should emphasize what is new and should not refer to purported merits or speculative applications of the invention as in lines 1-3. MPEP § 608.01(b) contains examples.

4. The Examiner notes, without objection, the possibility of informalities in the specification. The specification has not been checked to the extent necessary to determine the presence of all possible minor errors. The Applicant's cooperation is requested to consider correcting minor errors of which the Applicant may become aware during normal review and revision of the disclosure.

At page 2, line 4, should the word "inforamtiion" be --information--?

Claim Informalities

5. Claims 10-12 and 17-20 are objected to as being (directly or indirectly) dependent upon a rejected base claim. See MPEP § 608.01(n)V. The claim(s) would be allowable over the prior art of record if rewritten to include all of the limitations of the base claim and any intervening claims.

6. Claim 15 is objected to under 37 CFR 1.75(a) because the meaning of the phrase "the longer sequence" (page 32, line 1) needs clarification. Because no longer sequence was previously recited, it may be unclear as to what element this phrase refers. To further timely prosecution and evaluate prior art, the Examiner has interpreted this phase to refer to --a longer sequence--.

7. Claim 16 is objected to for the same reasons as claim 15 because the limitations are recited using obviously similar phrases.

8. The Examiner notes, without objection, the possibility of informalities in the claims. The Applicant may wish to consider changes during normal review and revision of the disclosure.

In claim 1 (page 28, line 17), should the word "which" be --which--?

Claim Rejections - 35 USC § 102

9. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Nomura

10. Claims 5, 7, and 14 are rejected under 35 U.S.C. 102(b) as being anticipated by Nomura [US Patent 4,862,504].

11. Regarding claim 14, Nomura [at abstract] describes speech processing recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

a contiguous sequence of phonemes without clear distinction containing at least one vowel [at column 2, lines 59-67, as a series of phonemic symbols in a respective syllable with a vowel];

treating it as Extended CV that is a unit that can not be split any more [at column 2, lines 58-60, as assume that the syllable CV combination is defined as the unit of speech synthesis].

12. Regarding claim 7, Nomura [at abstract] describes speech processing recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

Extended CV refers to a contiguous sequence of phonemes without clear distinction containing at least one vowel [at column 2, lines 59-67, as the syllable CV combination is defined as the unit of speech synthesis for a series of phonemic symbols in a respective syllable with a vowel];

it is provided with division information per Extended CV [at column 2, line 67-column 3, line 1, as respective syllable parameters are obtained];

phonetic information data is handled in it [at column 2, lines 66-67, as phonemic series are divided into syllables];

the phonetic information data is used for speech processing [at column 3, lines 41-42, as a series of speech parameters are generated for the series of phonemic symbols];

a computer readable storing medium storing the phonetic information data [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables are read out by the CPU].

13. Regarding claim 5, Nomura [at abstract] describes a dividing process recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

Extended CV refers to a contiguous sequence of phonemes without clear distinction containing at least one vowel [at column 2, lines 59-67, as the syllable CV combination is

defined as the unit of speech synthesis for a series of phonemic symbols in a respective syllable with a vowel];

a step where phonetic information is divided into it [at column 2, lines 66-67, as phonemic series are divided into syllables];

the phonetic information is received [at column 2, lines 61-64, as data is supplied and a series of phonemic symbols is given];

a computer readable storing medium storing a program for executing the process using a computer [at column 3, lines 48-57, as a program memory for the generator with a CPU].

Claim Rejections - 35 USC § 103

14. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

Nomura and Tenpaku

15. Claims 1-4, 6, 8, 9, 15, and 16 rejected under 35 U.S.C. 103(a) as being unpatentable over Nomura [US Patent 4,862,504] in view of Tenpaku [US Patent 6,317,713], already of record.

16. Regarding claim 8, Nomura [at abstract] describes speech processing recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

Extended CV refers to a contiguous sequence of phonemes without clear distinction containing at least one vowel [at column 2, lines 59-67, as the syllable CV combination is defined as the unit of speech synthesis for a series of phonemic symbols in a respective syllable with a vowel];

phonetic information is in it [at column 2, lines 66-67, as phonemic series are divided into syllables];

a phoneme dictionary contains a function of each phoneme associated with the information [at column 3, lines 48-57, as parameter files constitute the series of phonemic symbols in units of syllables];

the function is a contour [at column 5, lines 4-5, as parameter data is a level distribution plot];

the function is a vocal tract function [at column 2, lines 41-42, as the parameters represent tracheal characteristics];

a computer readable storing medium storing the phoneme dictionary [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables are read out by the CPU].

Nomura does not explicitly describe that the vocal tract function is a vocal tract transmission function; Nomura [at column 5, lines 4-12] uses spectral level as an example.

Tenpaku [at column 1, lines 12-17] describes speech synthesis, points out that human speech utterance is known by artisans to involve articulation by the vocal tract, and that modeling speech by waveform data simplifies the representation. Tenpaku also describes:

a contour of vocal tract transmission function of each phoneme [at column 12, lines 31-44, as time curves of vocal tract transmission characteristic, which are generated in accordance with the phoneme symbol string];

Tenpaku [at column 11, lines 24-28] points out that a wave representation allows rhythm control for synthesized speech and allows synthesis of other dialects and languages. It would have been obvious to one of ordinary skill in the art of speech synthesis at the time of invention to include Tenpaku's concept of filtering a waveform representation of speech by a vocal tract transmission filter characteristic of the phoneme being synthesized with Nomura's speech synthesizer and its database of speech parameters because that would provide the advantage of prosody control for more natural sounding synthesized speech.

17. Regarding claim 6, Nomura [at Fig. 1, items 3a-3d] describes a speech database recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

Extended CV refers to a contiguous sequence of phonemes without clear distinction containing at least one vowel [at column 2, lines 59-67, as the syllable CV combination is defined as the unit of speech synthesis for a series of phonemic symbols in a respective syllable with a vowel];

sample speech data is in it [at column 2, lines 66-67, as phonemic series are divided into syllables];

a data area that stores the sample speech data divided into it [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables];

a phonetic information area that stores phonetic information associated with the sample speech data [at column 3, lines 48-57, as parameter files constitute the series of phonemic symbols in units of syllables];

a computer readable storing medium storing a speech database [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables that are read out by the CPU].

Nomura does not explicitly describe that the sample speech data is waveform data; Nomura [at column 5, lines 4-12] uses spectral level as an example.

Tenpaku [at column 1, lines 12-17] describes speech synthesis, points out that human speech utterance is known by artisans to involve articulation by the vocal tract, and also describes:

speech waveform data [at column 1, lines 12-17].

Tenpaku [at column 11, lines 24-28] points out that a wave representation allows rhythm control for synthesized speech and allows synthesis of other dialects and languages. It would have been obvious to one of ordinary skill in the art of speech synthesis at the time of invention to include Tenpaku's concept of waveform representation of speech with Nomura's speech synthesizer and its database of speech parameters because that would provide the advantage of prosody control for more natural sounding synthesized speech.

18. Regarding claim 4, Nomura [at title] describes speech synthesis recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

Extended CV refers to a contiguous sequence of phonemes without clear distinction containing at least one vowel [at column 2, lines 59-67, as the syllable CV combination is defined as the unit of speech synthesis for a series of phonemic symbols in a respective syllable with a vowel];

generating speech data in it [at column 2, lines 66-67, as phonemic series are divided into syllables];

obtaining speech data in it [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables are read out];

the speech data is to be composed by concatenating the speech data in it [at column 5, lines 12-27, as lag and advance the speech parameters curves and sequentially interpolate to supply a series of speech parameters];

a step where phonetic information is divided into it [at column 2, lines 66-67, as phonemic series are divided into syllables];

the phonetic information is received [at column 2, lines 61-64, as data is supplied and a series of phonemic symbols is given];

the phonetic information is of a speech sound to be synthesized [at column 2, lines 54-55, as the series of parameters are supplied to a speech synthesizer];

a computer readable storing medium storing a program for executing the process using a computer [at column 3, lines 48-57, as a program memory for the generator with a CPU].

Nomura does not explicitly describe that the sample speech data is waveform data;
Nomura [at column 5, lines 4-12] uses spectral level as an example.

Tenpaku [at column 1, lines 12-17] describes speech synthesis, points out that human speech utterance is known by artisans to involve articulation by the vocal tract, and also describes:

speech waveform data [at column 1, lines 12-17].

Tenpaku [at column 11, lines 24-28] points out that a wave representation allows rhythm control for synthesized speech and allows synthesis of other dialects and languages. It would have been obvious to one of ordinary skill in the art of speech synthesis at the time of invention to include Tenpaku's concept of waveform representation of speech with Nomura's speech synthesizer and its database of speech parameters because that would provide the advantage of prosody control for more natural sounding synthesized speech.

19. Claim 3 sets forth limitations similar to limitations set forth in claim 4. Nomura and Tenpaku describe and make obvious the limitations as indicated there. Nomura also describes additional limitations as follows:

converting the provided speech data into analog signals of speech sound [at column 2, lines 54-57, as synthesizer generates speech from supplied speech parameters];

means for dividing, means for composing, and analog converting means [at column 3, lines 48-57, as a program memory for the generator with a CPU].

20. Regarding claim 15, Nomura also describes:

as a vowel element, a vowel (or others) [at column 2, lines 59-67, as the syllable CV combination with a vowel];

and a longer sequence shall be first selected as Extended CV [at column 2, lines 59-67, as a series of phonemic symbols in the syllable CV combination is defined as the unit of speech synthesis for a respective syllable].

21. Claim 2 sets forth limitations similar to limitations set forth in claim 4. Nomura and Tenpaku describe and make obvious the limitations as indicated there. Nomura also describes additional limitations as follows:

the speech data is obtained from a speech database [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables are read out of parameter files];

the speech data in the database is constructed with sample speech data [at column 3, lines 48-57, as parameter files constitute the series of phonemic symbols in units of syllables].

22. Claim 16 sets forth additional limitations similar to limitations set forth in claim 15. Nomura describes the additional limitations as indicated there.

23. Regarding claim 1, Nomura [at title] describes speech synthesis recognizable as a whole to one versed in the art by explicitly describing the content and functionality of the recited limitations as the following terminology:

speech database storing means storing a speech database created by dividing sample speech data obtained from recording human speech utterances into speech units [at column 3, lines 2-40, as files prepared beforehand storing natural speech parameters in units of syllables];

and associating the sample data in each speech unit with their corresponding phonetic information [at column 3, lines 41-42, as a series of speech parameters are generated for the series of phonemic symbols];

speech waveform composing means for dividing phonetic information into speech units [at column 2, lines 54-67, as series of speech parameters generator divides the phonemic series into syllables];

the phonetic information is received [at column 2, lines 61-64, as data is supplied and a series of phonemic symbols is given];

the phonetic information is of a speech sound to be synthesized [at column 2, lines 54-55, as the series of parameters are supplied to a speech synthesizer];

and for obtaining that (phonetic-corresponding, speech unit) speech sample data from the speech database [at column 3, lines 48-57, as parameter files of the series of phonemic symbols in units of syllables are read out];

and for generating speech data to be composed by concatenating that (sample, speech-unit) speech data [at column 5, lines 12-27, as lag and advance the speech parameters curves and sequentially interpolate to supply a series of speech parameters];

analog converting means converting that (composed) speech data into analog signals [at column 2, lines 54-57, as synthesizer generates speech from supplied speech parameters];

the sample speech data for storing is divided into speech units of Extended CV, which is a contiguous sequence of phonemes without clear distinction containing a vowel (or some

vowels) [at column 2, lines 59-67, as the syllable CV combination is defined as the unit of speech synthesis for a series of phonemic symbols in a respective syllable with a vowel].

Nomura does not explicitly describe that the sample speech data is waveform data; Nomura [at column 5, lines 4-12] uses spectral level as an example.

Tenpaku [at column 1, lines 12-17] describes speech synthesis, points out that human speech utterance is known by artisans to involve articulation by the vocal tract, and also describes:

speech waveform data [at column 1, lines 12-17].

Tenpaku [at column 11, lines 24-28] points out that a wave representation allows rhythm control for synthesized speech and allows synthesis of other dialects and languages. It would have been obvious to one of ordinary skill in the art of speech synthesis at the time of invention to include Tenpaku's concept of waveform representation of speech with Nomura's speech synthesizer and its database of speech parameters because that would provide the advantage of prosody control for more natural sounding synthesized speech.

24. Claim 9 sets forth additional limitations similar to limitations set forth in claim 15.

Nomura describes the additional limitations as indicated there.

Nomura and Tenpaku and Nishimura

25. Claims 13 and 21 rejected under 35 U.S.C. 103(a) as being unpatentable over Nomura [US Patent 4,862,504] in view of Tenpaku [US Patent 6,317,713], already of record, and Nishimura et al. [European Patent Application EP 0 821 344 A2].

26. Regarding claim 13, Nomura and Tenpaku retrieve speech parameters from a database. However, Nomura and Tenpaku consider all of the retrieved speech units to select for best providing natural synthesized speech. Neither Nomura nor Tenpaku explicitly describes searching for the speech unit by decreasing length.

Nishimura [at abstract] also describes speech synthesis using units selected from a speech waveform database. Nishimura also describes:

searching in order of decreasing length [at column 1, lines 52-55, as apply the longest matching method];

the string representing the reading is kana character [at column 3, lines 3-5, as transcriptions are accumulated in the waveform database as kana characters].

Nishimura [at column 1, lines 44-55] points out that it is possible to select a speech piece so that high quality sound can be synthesized without complex calculations. Following Nishimura, it would have been obvious to one of ordinary skill in the art of speech synthesis at the time of invention to include the concept of searching a database of kana in the order of decreasing length with Nomura and Tenpaku because the longest suitable unit will provide quality speech with no additional processing, or at least less intensive processing.

27. Claim 21 sets forth additional limitations similar to limitations set forth in claim 13. Nishimura describes the additional limitations as indicated there.

Conclusion

28. The following references here made of record are considered pertinent to applicant's disclosure:

Kandefer et al. [US Patent 5,153,913] describes extracting PCM speech parameters in coarticulated units from recorded utterances.

Hutchins [US Patent 5,384,893] describes prosody enhancement for synthesis of speech using syllable stress to modify parameters retrieved from a large database of speech waveform units.

Hasegawa [US Patent 5,463,713] describes speech synthesis following an accent pattern chosen according to vowel mora, vocal consonant+vowel, or voiceless consonant+vowel.

Saito et al. [US Patent 5,715,368] points out that a database of CV and VC units for speech synthesis has been widely used.

Arai et al. [US Patent 5,950,152] describes text-to-speech synthesis by chains of connected waveforms selected from a large database.

Nishimura et al. [US Patent 6,035,272] describes the same as EP 0 821 344 A2.

29. Any response to this action should be mailed to:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

or faxed to:

(703) 872-9306, (for formal communications intended for entry)

Or:


(703) 872-9306, (for informal or draft communications, and please label "PROPOSED" or "DRAFT")

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington, VA (Sixth Floor, Receptionist)

30. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Donald L. Storm, of Art Unit 2654, whose telephone number is (703) 305-3941. The examiner can normally be reached on weekdays between 8:00 AM and 4:30 PM Eastern Time. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703) 305-9645.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

April 5, 2004


Donald L. Storm
Patent Examiner
Art Unit 2654